Univerzita Karlova v Praze

Filozofická fakulta

Ústav anglického jazyka a didaktiky

Bakalářská práce

The Ability of Czech Speakers to Imitate the Temporal Structure of English Words with and without Czech Analogues

Schopnost českých mluvčích imitovat temporální strukturu anglických slov, která mají či nemají české analogy

Jan Jakšič

2015

Vedoucí práce: Mgr. Pavel Šturm

Acknowledgements:

I would like to express my acknowledgements to Pavel Šturm for his encouragement, comments, suggestions, and patience. It was a pleasure for me to have him as the supervisor of my thesis. Furthermore, I would like to thank Sarah Jane Gráfová, for her willingness to record the native models of the analysed words, and the Czech speakers, who provided us with the material for the research. Finally, my thanks go to Kateřina Brabcová, who encouraged me in the course of the thesis elaboration, and to all my other friends and my family, whose support helps me to overcome any problems.

Poděkování:

Rád bych vyjádřil své poděkování Pavlu Šturmovi za jeho podporu, komentáře, návrhy a také trpělivost. Byla pro mne radost spolupracovat s ním jako s vedoucím této práce. Dále bych rád poděkoval Sarah Jane Gráfové za její ochotu nahrát modelové realizace anglických slov a českým mluvčím, jejichž nahrávky posloužily jako materiál pro náš výzkum. Mé díky patří také Kateřině Brabcové, která mne povzbuzovala v průběhu celé práce, a rovněž i mým ostatním přátelům a členům mé rodiny, jejichž podpora mi pomáhá překonat jakékoli problémy.

PROHLÁŠENÍ

Prohlašuji, že jsem bakalářskou práci vypracoval samostatně, že jsem řádně citoval všechny použité prameny a literaturu a že práce nebyla využita v rámci jiného vysokoškolského studia či k získání jiného nebo stejného titulu.

V Praze dne 10. 8. 2015 …….

Souhlasím se zapůjčením bakalářské práce ke studijním účelům.

I have no objections to the BA thesis being borrowed and used for study purposes.

ABSTRACT

The aim of the current thesis is to ascertain whether there is an effect of Czech analogues on the temporal structures of English words in realization of Czech speakers. The theoretical part of the thesis introduces three major areas of our focus: foreign accent, word stress, and duration. Definitions of relevant terms are presented and existing findings about the areas are summarized in this part. The empirical part consists of a research; the Czech speakers imitate English words with and without analogues after listening to a native speaker's realization. The temporal structures in realizations of the Czech speakers are compared to those in the native realizations. Apart from examining the temporal structures with respect to analogues, we also try to determine the possible effect of Czech stress placement on the durations of the English words and their segments in Czech realizations. The acquired data are analysed with regard to their significance and the results are subject to a discussion.

Key words: Foreign Accent, Duration, Lexical Stress, the English of Czech Speakers, Analogues

ABSTRAKT

Cílem této bakalářské práce je zjistit případný vliv českých analogů na temporální strukturu anglických slov realizovaných českými mluvčími. Teoretická část práce představuje tři hlavní oblasti, na které se soustředíme: cizinecký přízvuk, slovní přízvuk a trvání. Jsou zde obsaženy definice příslušných termínů a shrnutí dosavadních poznatků z výše zmíněných oblastí. Praktická část obsahuje samotný výzkum. Čeští mluvčí imitují anglická slova s analogy a bez nich poté, co tato slova slyšeli v realizaci rodilé mluvčí. Temporální struktury realizací českých mluvčích jsou porovnány s temporálními strukturami v realizaci rodilé mluvčí. Kromě zkoumání temporálních struktur s ohledem na přítomnost analogů se snažíme zjistit případný vliv slovního přízvuku v češtině na trvání anglických slov a jejich segmentů v podání českých mluvčích. Získaná data jsou analyzována z hlediska jejich významnosti a výsledky jsou podrobeny diskuzi.

Klíčová slova: cizinecký přízvuk, trvání, slovní přízvuk, angličtina českých mluvčích, analogy

LIST OF TABLES AND FIGURES

TABLE OF CONTENTS

1 INTRODUCTION

As contact with people from different linguistic environments becomes a frequent matter in today's world, studying second language (L2) acquisition gains importance more and more. In the observations and analyses of L2 production, the phenomenon of foreign accent is a crucial one and any deeper knowledge in this area may contribute to elimination of unwanted effects in communication among people from different linguistic environments. When comparing two languages, some features of these languages differ more than others, and it is the interference of first language (L1) and second language in these features what is expected to be the trigger of a perceived foreign accent in the non-native speech.

The current thesis is focused on the ability of Czech speakers to imitate the temporal structures of separate English words in the production of a native English speaker. Studies (Tajima et al, 2007; Skarnitzl, 2005) have shown that duration of speech segments influences the perception of speech and the temporal distinctions between Czech (here L1) and English (here L2) may also contribute to the emergence of foreign accent. Another area in which Czech and English phonological systems significantly differ is word stress. However, it has not been sufficiently clarified which of the features of word stress is the most important in the indication of foreign accent in the English of Czech speakers. The phenomenon of speech segments duration is thus approached also from the viewpoint of segmental prominence, i.e. as one of the acoustic correlates of word stress. In these aspects, we try to find out whether the temporal structures in the L2 production of Czech speakers differ systematically from that of a native speaker.

A special emphasis is placed on the potential interference of Czech lexicon in the production of separate English words. Many Czech and English words have the same origin but their sound structure has been adjusted in accordance with the phonological system of the respective language. Whether the presence and relative frequency of individual English words in Czech lexicon – after the pronunciation adjustment – influence the production of such words by Czech speakers is one of the main tasks of the current thesis. From this viewpoint the thesis follows to a certain extent the diploma thesis of Jan Růžek (2013).

The thesis is divided into a theoretical part and an empirical part. The theoretical part explains the basic terminology and summarizes the existing findings about foreign accents and the phenomena observed in particular: word stress and duration. The empirical part introduces the hypotheses, the material and the method, and the results of the research. It also proposes conclusions based on the results and suggestions for further research. The thesis also includes the used bibliography and an appendix.

2 THEORETICAL BACKGROUND

2.1 Foreign accent

Knowledge of a second language in today's world is not only a very common thing but almost a necessity or at least a highly recommended quality. Understandably, acquisition of an L2 brings along many problems, which can appear on all levels of language – pragmatic, semantic (e.g. idiomatic), grammatical (syntactic and morphological) and phonological. The difficulties which the speakers of a different first language have on the last mentioned level lead to what is generally called foreign accent.

The term "foreign accent" has been subject to many definitions without any successful consensus. Moyer (2013) deals with differentiation of accent from a) dialect – comprising also grammar, vocabulary and discursive style, and b) pronunciation – referring exclusively to articulation. Her definition of accent is:

> "Accent is a set of dynamic segmental and suprasegmental habits that convey linguistic meaning along with social and situational affiliation." (Moyer, 2013: 11)

Moyer applies this definition also to a non-native variety despites being aware of the differences in the abilities of L2 speakers to "convey all the nuances" (Moyer, 2013: 12). Another definition is proposed by Roy Major:

> "Mastering the phonology of a language involves (a) individual segments ..., (b) combinations of segments, which produce syllables ..., (c) prosody (stress, rhythm, tone, intonation), and (d) global accent, or the overall accent of a speaker. A global foreign accent is the result of non-native combination of (a), (b), and (c). If one masters native-like pronunciation of one or two but not all three levels, then a foreign or non-native accent results." (Major 2001: 12)

Whether there is or there is not an exact definition of the term, we broadly understand it as the sum of divergences from the native norm in the speech production of a non-native speaker.

What exactly leads to the presence of foreign accent in one's speech and which features of pronunciation under which circumstances contribute to the perception of a foreign accent has been studied by linguistics for decades. The closer cross-linguistic contact in today's world, the differences in the abilities of non-native speakers to acquire a foreign language and also the characteristics of foreign accent perception have led to the increase of interest in L2 acquisition, its potential limits and the processes aligned to it. Studying foreign accents can contribute to the area of language learning in many ways, as affirmed by Piske et al (2001: 192): "[Studying perceived foreign accent in the speech of L2 learners] may help resolve some theoretical issues regarding whether there are age-based constraints on L2 learning. In addition, identifying factors that influence degree of L2 foreign accent may be important for the teaching of second languages." The environments and situations, listed by Volín and Skarnitzl (2010), show other effects of accents on our lives and the importance of their observations.

Although there may be some possibilities of positive use of foreign accents (e.g. in drama), receiving negative reactions because of foreign accent is more likely. The stigmatization of accents may lead to discrimination or at least to lower estimation of L2 speakers. Studies of Brennan and Brennan (1981) or Rubin (1992) (both cited in Volín and Skarnitzl, 2010) showed the tendency of listeners to base the evaluation of L2 speakers on the presence of foreign accent in their speech. Volín and Skarnitzl (2010: 271) conclude that "we should take into account the possibility that foreign accentedness may have a bearing on the quality of people's lives and it should be studied seriously and thoroughly."

In consequence of such prejudice and along with the globalization process, the desire of L2 speakers to achieve at least a certain level of native-like pronunciation is expected to be increasing. This aim is, however, not as easily reachable as it might seem, because, as Volín and Skarnitzl (2010: 271) assert, "our current knowledge is too imperfect to design exercises which guarantee fast progress to everyone." A continuous study of foreign accents will thus be necessary if we want to broaden the possibilities of overcoming problems arising from them.

2.1.1 Features of foreign accent

L2 learners have problems with pronunciation of segments and sequences that are absent in their native language and it is believed to be so because of a "filter" of the L1

phonological characteristics (Strange, 1999). Several models exist that describe the processes during L2 acquisition – Native Language Magnet Model (Kuhl and Iverson, 1995), Perceptual Assimilation Model (Best, 1995) or Speech Learning Model. The last of these was developed by Flege (1995) and is to be described fully in the following paragraphs. To understand the issue of L1 and L2 interference, we must also understand the different levels on which the various aspects of foreign accent are manifested.

Levels of foreign accent manifestation

Strange (1999: 2518) asserts that apart from "learning to perceptually differentiate phonetic segments distinctive in the new language, but not in the first language […] L2 learners must also learn the language-specific allophonic, phonotactic and prosodic constraints." Many studies have tried to determine which of these levels prevail in the detection of foreign accent by native listeners.

Braun et al (2011) in their study of Dutch pronunciation of English found that in studying word stress, the vowel quality (i.e. the segmental level) contributes much more to the perception of accentedness. Other studies (Flege and Hillenbrand, 1984; Tajima et al, 1997) show that the distribution of the markers of foreign accent between the levels may be equal. Moreover, Piske and colleagues (2001: 212) state that "when asked, native speakers usually report that they perceive non-native speech as foreign-accented because of both segmental as well as suprasegmental errors produced by bilinguals" and Flege and colleagues (1995: 3132) also claim that previous studies proved foreign accent to be "cued by segmental, subsegmental, and prosodic divergences from the phonetic norms of English."

Segmental level of foreign accent manifestation

The segmental level of speech comprises the characteristics of individual speech segments, including their allophonic realization. Therefore, the most important features that contribute to the deviations from native norms of pronunciation are the quality and the duration of the segments.

The role of L1 in the processing of "unfamiliar" L2 segments has been repeatedly proved by empirical research (Flege and Hillenbrand, 1984; Flege et al, 1995; Strange, 1999). Leather and James (1996: 273-4) argue that a "beginning learner

seeking to impose phonetic structure on the L2 speech to which he is exposed makes perceptual reference to the phonetic categories of his Ll." It was Flege (1995) with his Speech Learning Model (SLM) who contributed significantly to the studies of interference of L1 and L2 segments.

SLM works with a continuum of L2 sounds comprising segments almost identical with L1 segments on one extreme of the scale, segments dissimilar with any L1 category on the other extreme, and all segments in between. The L2 speech sounds which are similar to a certain native category might be ascribed to such a category, resulting in a change in pronunciation. This sound may assimilate completely to the L1 category, or it may be pronounced in an intermediate fashion, differing from the native categories in both L1 and L2. Flege's example is on the establishment of L2 categories in English of Spanish speakers: "If Spanish learners of English are unable to establish a category for English /æ/, the model predicts that they will produce English /æ/ with Spanish /a/-like properties, and vice versa. (That is, their L2 /æ/ will have F2 (second formant) values that are too low, and their L1 /a/ will have F2 values that are too high.)" (Flege, 1995: 243-4).

On the other hand, the L2 segments which are significantly distinct from any L1 sounds might base their own position in the phonetic inventory of an individual and thus be produced more accurately in terms of L2 native speaker's production. Once more in Flege's words, that means that "a native Spanish speaker should be more likely to establish a phonetic category for English /æ/ or /ə/ than for English /i/ (which differs only slightly from Spanish /i/)" (Flege, 1995: 243).

The validity of this approach has been proved also in other studies. In the general discussion of their study of limits on pronunciation accuracy in adult foreign language speech production (1984: 717-8), Flege and Hillenbrand state:

> "The first conclusion to be drawn from this study is that adult native speakers of English may produce new phones in a foreign language (such as French /y/) more accurately than L2 phones which have a clear counterpart in the native language (such as French /u/). Listeners' identification of vowels in experiment 1 revealed a tendency for inexperienced American speakers of French to produce the new vowel /y/ more accurately than /u/. This suggested that new L2 phones

> may be learned more rapidly than L2 phones which have a clear counterpart in L1."

In the same way, Flege's SLM is also applicable for consonants (Flege, 1995: 254-62).

Suprasegmental level of foreign accent manifestation

In longer sequences and mainly in connected speech, foreign accent emerges and is perceived also on the suprasegmental level of speech. Nevertheless, Flege's SLM was designed for the segmental level of speech. The listeners in Piske et al (2001) reported to have perceived influence of both segmental and suprasegmental elements of non-native speakers' pronunciation. In addition, Piske et al state that "segmental and suprasegmental aspects of speech are very closely related, so that in many cases it is difficult to draw a clear distinction between the two" (Piske et al, 2001: 212). This was evidenced also by Braun et al (2011) whose studies have shown effect of not only vowel quality (segmental level) but also intonation (prosodic level – Braun et al 2011b, cited in Braun et al 2011).

Since the majority of researches focused on and proved an effect of deviation from L2 segmental features on accentedness, Volín and Skarnitzl (2010) studied the possible influence of suprasegmental features of Czech English on the perception of foreign accent. The results, proving the effect of prosodic parameters on the perception of accented speech, are yet another evidence for claims that accentedness can be ascribable also to suprasegmental errors in speech.

Even though Flege's SLM is attested mainly on the segmental level of L2 acquisition, its rules and predictions could be extended also to the suprasegmental level. If the L1 provides us with sufficiently distinguishable categories not only on the level of segments (e.g. positions of vowels) but also on the suprasegmental level (e.g. presence or absence of a word in the L1 lexicon), it is plausible to expect the applicability of SLM in a similar way. The possibility of such a focus shift was demonstrated also by McAllister et al (2002) and Mennen (1999; cited in Piske, 2001).

Application of SLM on duration and presence or absence of analogues in L1

Application of SLM on both segmental and suprasegmental level is possible also in studying duration of English vowels in the pronunciation of Czech speakers. Czech

phonological system recognizes long and short vowels. With only five vowels distinguished by vowel quality, duration is a major factor in word structure and long and short vowels function as minimal pairs (cf. *vir* (virus) and *vír* (whirl)[1]). On the other hand, English has twelve vowels and eight diphthongs and vowel duration plays a subsidiary role in English prosody, which is utilized in the realization of stress and vowel quality (i.e. the presence or absence of vowel reduction.)

One of the major aims of the current thesis is to find out whether there is also an influence of presence of L2 words in L1 lexicon (adjusted to the L1 phonological rules) on the production of such words in L2. As already mentioned, this thesis continues in the steps of the diploma thesis of Jan Růžek (2013). Růžek analysed the segmental duration of English words in the production of Czech speakers. The results showed a tendency of Czech speakers to omit vowel reduction and not to exaggerate stress by the means of duration. In two-syllable words with stress on the initial syllable, this syllable was usually shorter and the second syllable longer than in the realization of a native speaker. In words where the Czech analogue includes a long vowel or has a similar structure as its English counterpart, the differences were smaller. In words with stress in the final position, the stressed syllable was usually shorter than in the native realization. However, there was not any compensation observed on the first syllable.

Since Růžek included only words with Czech analogues, it is not clear, whether his results can be ascribed to the presence of analogues. Involvement of words with analogues (A+) and also of words without analogues (A–) would help to answer the question. We have shown the possibility of application of SLM on features different from segments quality. Therefore, it could be expected that, as a consequence of the interference of Czech and English, the word and segments durations in the words with Czech analogues would be less accurate than in the words without them.

2.1.2 Factors influencing the degree of foreign accent

Emergence of foreign accent is subject to various factors and their combinations. Although a foreign accent is on the one hand a display of articulatory deviations from native speech production standard, the presence or the absence and especially

[1] A recent study on the spectral quality of Czech vowels (Skarnitzl and Volín, 2012) suggests that the high front vowel moved towards central position when short.

the strength of a foreign accent in L2 speakers' pronunciation is always a matter of abstract estimation by native speakers. With respect to this, the factors affecting foreign accent cannot be only those linked to the speakers, but must comprise the information about those who perceive the strength of the accent. Also the procedures chosen by the researchers play an important role in the way a particular speech unit is judged by the control group. In the same way, we approach the variety of accent-affecting factors.

Any general observations about the factors which might have any persisting effect on foreign accents are made difficult by the similarities and differences between the studied L1 and L2. The formal relationship between the two target languages is displayed on both segmental level (the phonological inventories of the languages) and suprasegmental level (e.g. rhythmical parameters of the languages), which are discussed in the previous chapter.

Factors on the side of the speakers

The factors linked to L2 speakers are undeniably the most numerous when approaching accents. There were various attempts to provide a general view on these factors on the basis of previous research (Major, 1987; Long, 1990; Thompson, 1991 – all cited in Piske, 2001). Leather and James (1996), for example, distinguish maturational constraints on one hand and individual and social constraints on the other. They focus on the individual and social constraints and include among them the factors of motivation (both materialistically and socially triggered), oral and auditory capacities, psychological characteristics or gender (also in Flege, 1996). Other constraints having an effect on the presence or absence of the foreign accent are the Length of Residence (LOR) and especially the Age of Learning (AOL) (Flege et al, 1995; Piske, 2001). In this subchapter, we will discuss these two and some other factors especially in connection with studies of Flege et al (1995) and Piske (2001).

Studying AOL contributes to the assignation of the critical periods in L2 learning. These periods mean:

a) the lowest possible age of the beginning of L2 acquisition, after which some traces of foreign accent would be detected by native listeners;

b) the highest possible age of the beginning of L2 acquisition, at which no foreign accent appears

Flege and colleagues studied native English speakers' perception of accentedness in speech of Italians living in Canada for more than 15 years with different AOL. The study suggests that a foreign accent of L2 can be perceived in speech of people who began to learn L2 as soon as when 3.1 years old, which is almost twice sooner than in conclusions of Long (1990) (cited in Flege, 1995). 11.6 years was the other extreme (the mean critical period being at 7.4yo), which on the other hand corresponded quite well with Long (12yo). No Italian living in Canada who started to learn English after the age of fifteen was labelled as without accent.

Another study dedicated to the determination of AOL influence was that of Piske. The results show that "AOL is the single most important predictor of degree of L2 foreign accent" (Piske, 2001: 212) because no other factor (gender, LOR, or L1 use) showed significance independently on AOL. Nevertheless, combining AOL with other factors can contribute to more precise observations, since the "ultimate attainment in the pronunciation of an L2 is dependent on various factors, not just on the state of neurological development at the age of first intensive exposure to the L2" (ibid: 204).

Because of the fact that Flege and colleagues worked with L2 speakers who had been living in the country with L2 as a primary language for more than 15 years, "even a small LOR effect is of interest" (Flege et al, 1995: 3132) and shows that L2 pronunciation acquisition is not a process which ends after five or ten years. The results in Piske show an opposite direction, i.e. after certain time, the growth of LOR has a slowing effect on the improvement of L2 pronunciation. However, the exact time when the amelioration starts to weaken is not known.

Several studies have focused on the factor of L1 use in the production of L2 (Flege et al, 1997; Suter, 1976 – cited in Piske, 2001). It was Piske again who dedicated himself to this phenomenon. In his study, he suggests that the native language background can also have an effect on the ability to learn L2 (Piske, 2001). This study

of English-Italian early and late bilinguals shows that the strength of foreign accent increases as a function of L1 use. As stated above, L1 use influence on foreign accent is interesting in correlation with AOL and other factors. The correlation with AOL and gender is present also in the study of Flege et al (1995).

It has also been shown that the effects of L1 use do not necessarily need to be of a long-time period character (Sancier and Fowler, 1997; cited in Piske, 2001). Even an exposure of several months to either L1 or L2 may trigger a drift in a bilingual's pronunciation, perceivable by listeners of the L2 speakers. Also the *mode* (e.g. monolingual vs. bilingual mode) of current exposure to either of the languages seems to play a role in the perception of accent (Grosjean, 1997, 2000; cited in Piske, 2001). Studying L1 as a factor of accent production and perception definitely provides phoneticians with a wide platform for future research. However, the interconnection of this factor with other phenomena such as motivation, AOL or LOR should be always considered.

Two more constraints on the side of the speaker can be mentioned. While gender does not play any role in Piske's study, there is an interesting observation in Flege et al (1995: 3132): "Female subjects who began learning English as children were found to pronounce English somewhat better than did males matched for AOL, whereas males who began learning English in late adolescence pronounce English somewhat better than did their female cohorts." Similarly ambiguous are findings about the effect of motivation. While Suter (1976) and Willing (1988) – both cited in Leather and James (1996) – found the speakers' concern about the level of their L2 important, Leather and James (1996) and Piske (2001) have doubts about that mainly because of the limited number of occupations requiring native-like abilities in pronunciation.

Factors on the side of the research method

The fashion in which a particular research is conducted also contributes to the conclusions that can be drawn from the results. With the complexity required in empirical research, it is often difficult to maintain an absolute independence of the findings. Different methods are used in dependence on the purpose of the research and their authors might have to choose one method or the other despites the fact that both of them have some benefits and some drawbacks.

In perception studies, differences may result from the length of the perceived sequences (individual segments, words, short or long utterances) or their intelligibility (real words or nonsense sequences). The variability of the perceived units and the number of their repetition also matters. Piske (2001) lists some of the possible elicitation techniques: reading materials of different length, repetition after hearing a native speaker model, or production of free L2 speech. He discusses the validity of the latter technique stating that "subjects may be able to avoid difficult L2 sounds, sound sequences or even words if they are asked to talk about something freely" (Piske, 2001: 194) and prefers the use of repetition of fixed sets of materials (e.g. sentences). Klatt (1976), on the other hand, depreciates the method of uninterrupted repetition in durational researches because it permits "the subjects to build up a very stable psychological reference pattern against which to judge changes in duration" (Klatt, 1976: 1219). In terms of accent perception, an important question is also how it is possible to measure and rate the strength of a foreign accent. The study of Southwood and Flege (1999) shows that foreign accent actually is a continuum divisible in equal units and thus "it is appropriate to use an EAI [Equal Appearing Interval] scale in foreign accent studies" (Piske, 2001: 195).

Factors on the side of the listeners

The listeners' ability to judge foreign accents has not been studied as much as the previously mentioned factors. In spite of not including tests of these abilities in their study, Flege et al (1995) are aware of the effect this aspect can have on the results and speculates that there can be an influence of "the native dialects of Canadian English spoken by the ten listeners, idiolectal differences in their representations for segmental and prosodic characteristics of English, or differences in "tolerance" ranges for English phonetic structure" (Flege, 1995: 3132). In their study of perception of foreign accent in English of Dutch speakers, Nejjari et al (2012; cited in Růžek, 2013) found that listeners who come into contact with Dutch accent in English more often tend to be more tolerant in their evaluations. Piske (2001) suggests that there should not be any elimination of listeners and researchers should work with a representative sample of raters.

2.2 Description of selected phenomena in English and Czech

Since the current thesis is focused on the temporal structures of the English of Czech speakers, the major area of our focus is duration. We approach it from two perspectives.

a) duration as an acoustic correlate of word stress

b) duration as an individual feature influencing the production and perception of speech

These two perspectives are discussed in the following subchapters.

2.2.1 Stress and prominence

Studies which observed the perception of stress in the production of L2 speakers (Volín and Skarnitzl, 2010; Braun et al, 2011; Frost, 2011) show that stress significantly contributes to a detection of foreign accent. The term "stress" itself, however, can be problematic because it refers to various phenomena. The first major distinction must be made with regard to the level on which we focus. Stress can be used as a term describing a relative prominence of one syllable over another in an individual word. This is what we call the word (or lexical) stress. Stress appears also on the level above individual words where it signifies the relative prominence of a syllable or a word over other parts of an utterance. This is referred to as the prosodic stress.[2] Because of the character of this thesis, we work with stress on the lexical level, and thus the term "stress" is always used here in the first of the two meanings.

When dealing with stress, the basic division could be made on the stressed and the unstressed syllables which differ in that the stressed syllable "is somehow more prominent that the neighbouring unstressed syllables." (Skarnitzl, 2005: 183). This is, however, a simplified approach, since the matter is more complex in terms of English. Frost (2011: 68) points out that "there is some disagreement as to how many levels exist in English; for example, according to Pennington (...), between four and six levels suffice for a detailed transcription. Other authors, such as Cruttenden (...),

[2] Ladd (2008:49) states that "probably no topic in the general area of intonation and suprasegmentals has posed such a puzzle as stress". Some authors try to solve the terminological problem by using "stress" for word stress and "accent" for prosodic stress (Plag et al, 2011) and some speak about "accent" only in connection with intonation (Bolinger, 1958). Since this area is problematic, literature is not consistent in using these terms.

distinguish four levels: primary stress, secondary stress, tertiary stress and unstressed." Plag et al (2011: 362) speak about three basic types of syllables: primary syllables, secondary syllables and unstressed syllables.

Similarly to the current thesis, most studies about stress focused on the stressed and unstressed syllables interrelations. Nevertheless, some researchers have dealt with the distinct manifestations of primary and secondary stress. Plag and colleagues (2011) see the drawbacks of previous studies of stressed and unstressed syllables in that it is uncertain whether the results can be attributed to stress or to prosodic accentuation because the phenomenon of prosodic stress was not eliminated from these experiments. The results of their study show that F0, intensity and spectral quality were different in syllables with primary stress and with secondary stress. Surprisingly, such impact was not found in duration.

The difficulties with handling the term "stress" is captured by Frost (2011: 68), who states that "stress has always seemed to resist all attempts at definition: the closer one investigates the nature of stress, the more difficult it becomes to define." Its definition could be subject to lengthy discussions and still would hardly lead to a uniform conclusion. One of the main reasons for such difficulty is the necessity to approach stress from both speaker's and listener's point of view, i.e. to include production and perception in the analysis of stress.

An interesting description of stress as a local hyperarticulation is provided by de Jong (1995). Skarnitzl (2005: 183) summarizes general views on production of words stress by stating: "The production of a stressed syllable is typically said to involve greater muscular energy than the production of an unstressed syllable." In terms of words stress perception, the importance seems to rest on the prominence of the stressed syllable over the unstressed. The prominence is always relative, i.e. "the stressed syllable must be defined as prominent in relation to its surrounding environment." (Frost, 2011: 68). The current approach to prominence (as presented in Skarnitzl, 2005; Frost, 2011; Plag et al, 2011) identifies four acoustic correlates (clues) which contribute to the perception of a stressed syllable:

- pitch (F0)
- duration (quantity)
- intensity (loudness or amplitude)

- vowel quality (formant structure)

Stressed syllables in English tend "to have higher pitch, higher intensity and longer duration. Furthermore, only stressed syllables can contain the full range of vowel phonemes, while in unstressed position most vowel contrasts are absent" (Plag et al, 2011: 362). Languages differ in the extent to which they use these individual correlates in stressed syllables. Which of them prevails and which of them is less important in terms of English was subject to many studies (de Jong, 2004; Plag et al, 2011) whose results show ambiguous findings.

In his study (2011), Frost focused on English and French – two languages which differ significantly in terms of stress. His experiment involved three out of the four acoustic correlates of stress – F0, duration and vowel quality. Not only that the results show that the distinctions in word stress in the two languages are reflected also in their perception by native speakers but they also reveal the strength of the individual correlates: "The pitch contour which accompanies the prominent syllable is of great importance in English and, at least subjectively, the cue of pitch is relatively more important in English than in French, and perhaps the most important of all the cues" (Frost, 2011: 81).

Other studies show that different acoustic cues might make the crucial contribution to the perception of stress. The results of de Jong's study (2004) show a certain role of duration and also consonantal voicing. Mattys (2000; cited in Plag et al, 2011) found the importance of pitch, intensity and duration, which was confirmed by Plag et al (2011) in case of the first two cues. Despites the diversity in empirical findings, the once generally held opinion that intensity is the major acoustic cue in stress production and perception is now outdated. This is confirmed also on the level of prosodic stress (Bolinger, 1958). It seems more than probable that it is the combination of all of the cues that forms prominence.

Differences in stress distribution are present also in Czech and English. Stress in Czech is fixed on the first syllable of a stress foot and thus "the prominence of the first ("stressed") syllable is not (need not to be) too significant for perception" (Skarnitzl: 2005: 184). On the contrary, stress is free in English and may alter the meaning of words (e.g. noun /ˈriː meɪk/ vs. verb /riːˈmeɪk/) and it is, therefore,

expected to be of higher importance for speakers of English to mark the word stress properly.

The interrelations of Czech and English in terms of word stress have also been studied and they are discussed more in the subchapter The English of Czech speakers. For now, we can mention the study by Skarnitzl (2005) who focused on the word stress perception by Czech listeners. Despites the fact that "there was a significant correlation between the students' difficulties in recognizing the stressed syllable, expert assessment of the stressed syllable's prosodic prominence, and the objective salience of the stressed syllable" (Skarnitzl, 2005: 193), the results failed to show which of the acoustic cues contributed more and which less to the stress perception. In terms of production, the study of Volín and Weingartová (2014) shows that "Czech speakers realize English word stress differently from British speakers" (Volín and Weingartová, 2014: 181). Because duration, which is discussed closer in the following subchapter, has different functions in Czech and in English, we suppose that also this factor might be an important cue to stress realization and perception in the English of Czech speakers.

2.2.2 Duration

Unlike stress, duration is not a complex term and can be described as a temporal quantity of a specific unit. Duration is a phenomenon which can be manifested on both segmental and suprasegmental levels depending on the unit we focus on and its function in its environment. Both of these levels are of importance for the purpose of the current thesis. On the suprasegmental level, it is important to mention rhythmical classification of languages. In dependence on which of the units appears in more or less regular temporal intervals, we distinguish stress-timed, syllable-timed and mora-timed languages. English, as a stress-timed language, tends to shorten unstressed syllables so that they fit in a temporally delimited stress group. On the segmental level, duration is manifested in allophonic variations or in vowel reductions. In the current thesis, we deal with both of these two levels.

In the following paragraphs, the theoretical background is based mainly on Dennis H. Klatt's study: *Linguistic uses of segmental duration in English* (1976). He lists seven factors that influence the durational structure of a sentence: extralinguistic, discourse level, semantic, syntactic, word level, phonological/phonetic and physiological (Klatt, 1976: 1210). Some of these parameters (e.g. the physiological

factor) are common to English and Czech speakers and therefore have no significance for our purpose. As the current thesis works with separate words without any discourse or syntactic context, the most important factors for us will be the phonological/phonetic one.

In one respect, we could relate also to the semantic level. Klatt shows evidence that when words bear informational focus or they are new in a sentence, they tend to be of a greater duration than the same words in a context in which they do not have such a role. Also Umeda finds out "that semantic novelty has an influence on segmental durations in the sense that an unusual word is longest the first time that it appears in a continuous discourse (Umeda, 1975; cited in Klatt, 1976: 1210). It has to be emphasized that both observations are made on the level of conversation and from the perspective of the speaker. However, it is feasible to expect that – apart from a possible mispronunciation – the words that are unknown and thus new to listeners would be reproduced more slowly, at least for the first time.

Klatt (1976: 1213) distinguishes four different phonetic/phonological factors influencing durational structures. We will focus on three of them: inherent phonological duration, shortening of unstressed syllables, and influence of a postvocalic consonant on vowel duration.

Inherent phonological duration

The inherent phonological duration of segments shows great difference between English and Czech. In terms of vowels, the English vocalic inventory[3] is more numerous than the Czech one. The twelve spectral variations of English monophthongs are sufficient to distinguish individual phonemes and thus duration does not have a distinctive function. For instance, the word pair *fill* – *feel* (/fɪl/ – /fiːl/) differs not only in the duration of their vowels, but mainly in their spectral quality. Despites that, Klatt's previous studies found that „differences in duration as a function of segment type accounted for about half of the variance in stressed vowel durations in a connected discourse" (Klatt, 1976: 1213).

Unlike in English, duration has a distinctive function in the Czech vocalic system, since we distinguish only five types of vowels in terms of spectral quality. Czech short

[3] In the current thesis, all pronunciation references are made to RP [Received Pronunciation].

vowel /a/ and long vowel /aː/ are different phonemes because they function as a minimal pair in words *val* (wall, dike) and *vál* (past tense of "to blow"). The duration difference between short and long vowels in Czech has been studied by Skarnitzl (2012; cited in Skarnitzl and Volín, 2012). The results show that while Czech long /aː/ has 1.79 times longer duration than its short counterpart, the differences in high vowels are not that significant (/uː/:/u/ = 1.6; /iː/:/ɪ/ = 1.29). This is probably the reason why Czech short and long high front vowels differ also in their spectral quality (short /ɪ/ and long /iː/). We can therefore say that Czech has six vowel qualities rather than five. Moreover, the high back vowel also shows a tendency to move to central positions when short (Skarnitzl and Volín, 2012).

The inherent phonological duration applies also to English consonants: "Voiceless fricatives are about 40 ms longer in duration than the corresponding voiced fricatives. Small differences [...] are also observed as a function of place of articulation for consonants. Bilabial stops are typically slightly longer in duration than alveolars and velars." In Czech, voiced plosives are said to be shorter than voiceless plosives (Machač and Skarnitzl, 2007). It can be stated that duration of vowels and consonants in both languages thus always partly depends on their inherent characteristics independent of their environment. Klatt concludes that these "differences in inherent duration account for much of the variation in segmental timing in speech" (Klatt, 1976: 1213).

Shortening of unstressed syllables

Segmental duration as one of the acoustic cues of word stress was discussed in the previous chapter. As has been shown in many studies (Oller, 1973; Lehiste, 1975 – both cited in Klatt, 1976), stressed syllables tend to be longer in duration than unstressed syllables. In terms of vowel duration, the reduction of an unstressed vowel in English also contributes to the shortening of its duration. Klatt (1976: 1214) adds that "lexical stress also exerts to an influence on consonantal durations. Prestressed consonants are slightly longer in duration than other consonants, all else being equal." Stress in Czech is placed at the beginning of a stress unit, i.e. mostly a word, and therefore, when working with individual words, there are not any prestressed segments.

Influence of a postvocalic consonant on vowel duration

The duration of segments is very often influenced by its immediate environment. A typical example in English is a phenomenon usually called pre-fortis shortening. Cruttenden (2008, 95) asserts that "the length relationships between vowels are complicated by the influence of following voiceless consonants: /iː/ in *beat* is only about half as long as the /iː/ in *bee* or *bead* […]." In other words, English vowels are usually shorter before a voiceless consonant than when preceding its voiced counterpart in the same environment. Klatt uses words *bag* and *back* as an example. In a phrase final position, where this phenomenon is the strongest, the vowel /æ/ was shorter by 50 ms (approx. 20%) in *back* than in *bag* (Klatt, 1976: 1214).

Machač and Skarnitzl (2007) studied the effect of consonants on adjacent vowels. They tried to find out whether there was a vocalic compensation in Czech in dependence on the voicing of the adjacent consonant, i.e. whether "a voiceless (longer) consonant mean[s] a shorter adjacent vowel, and [...] a voiced (shorter) consonant mean[s] a longer vowel" (Machač and Skarnitzl, 2007: 538). The results suggest a vocalic compensation in adjacency of consonants, but the "compensation tendencies appear to be stronger in CV than in VC sequences" (Machač and Skarnitzl, 2007: 540). The percentage degree of the compensation in VC [vowel-consonant] sequences was 16% on average but only 3 out of 19 observed sequences were of statistical significance.

The observations discussed so far were focused primarily on the realization of speech and "segmental timing," which, as confirmed by Klatt, "carries a high functional load in English, providing information about the semantics, syntax, and the segmental composition of an utterance" (Klatt, 1976: 1208). Digressions from temporal structures in L2 production should thus have effect also on the perception of foreign accent and even influence intelligibility of utterances.

In their study (1997), Tajima and colleagues recorded both native and Chinese speakers of English. Their recordings were modified in such a manner that the native ones were warped to match the temporal structures of the non-native ones and vice versa. These modified recordings were mixed with the unmodified ones and played to native English listeners. The results showed that the 'corrupted' temporal structures

of the native speakers deteriorated the intelligibility by 5-15 percentage points, while the 'corrected' structures improved the intelligibility by 15-25 percentage points. Tajima and colleagues also mention several other studies (Jonasson and McAllister, 1972; Hutchinson, 1973 – both cited in Tajima et al, 2007) that show the tendency of listeners to value the speech of non-native speakers when the temporal structures are more similar to the native production.

Based on what we have stated, the influence of Czech (as L1) on the temporal structures of English (L2) words in the realization of Czech speakers can be hypothesized. This interference can happen either as a part of stress distribution within words (i.e., which syllable is stressed) but also independently from stress as a manifestation of inherent temporal characteristics of Czech phonological system (most notably distinctive vowel length).

2.3 The English of Czech speakers

Being a native language for approximately 10 million people, Czech is not one of the most usual L1 languages in foreign accent researches. Nevertheless, the less frequent a language is, the more important it is for its native speakers to learn some widely spoken language. English has been such a language for Czechs for more than two decades and studying Czech and English interrelations is thus a very important matter. Studies have been carried out to determine what makes English of Czech speakers deviated from its native spoken form. Researchers have focused both on the characteristics of production of English by Czech speakers and on the way Czech speakers perceive various features of native English.

The current thesis is in part based on the diploma thesis of Jan Růžek (2013) who dealt with the influence of Czech lexicon on temporal structures of the English of Czech speakers (see Foreign accent). Other studies have focused on differences in the perception and production of stress as this phenomenon differs significantly in the two languages. Skarnitzl (2005) studied English words stress in perception of Czech listeners and compared it to experts' assessment of the prominence and to the data from an acoustic analysis of the stressed syllables. The study did not determine the aspects of word stress which were mostly relied on by the students

or which contributed most to their errors. Nevertheless, the complexity of stress in perception of Czech listeners was observed as "the more conflicting acoustic cues there are concerning the stressed syllable, the greater difficulties students have with its identification. It seems that all the acoustic parameters analyzed in this study play some role in the perception of English word stress by Czech listeners" (Skarnitzl, 2005: 193).

Volín and Weingartová (2014) also dedicated themselves to stress; an incorrect prominence distribution might contribute to communication difficulties or even unintelligibility. They focused on the effect that the four acoustic correlates of stress (pitch, intensity, duration and spectral slope) have on the stress realization. Recordings from Czech and British female speakers were analysed and the results showed that for the Czech subjects the least native acoustic cue of stress was pitch, followed by intensity, and that also vocalic quality plays a role as the spectral differences between stressed and unstressed syllables were smaller in the Czech than in the British realization. Duration, on the hand, was relatively comparable to that of the native speakers. The overall observation was that Czech speakers tended to change the stress placement in accordance with the Czech system (i.e. towards the first syllable).

A suprasegmental level of foreign accent was observed by Volín and Skarnitzl (2010). One of the motivations for this focus was that "while certain vowel or consonant cues of foreign accent may be absent in an individual sentence whose accentedness is being assessed; there is always some rhythm, intonation and tempo present in every utterance" (Volín and Skarnitzl, 2010: 272). The research itself comprised the variables from time domain (articulation rate and speech rhythm), energy domain (loudness) and frequency domain (fundamental frequency – F0). The results showed that the suprasegmental aspects of speech reflect the strength of accentedness in Czech English. This was observed mainly for the frequency domain but also for the domains of time and energy.

There have been also other studies concerning English of Czech Speakers (e.g. Skarnitzl et al, 2005; Bissiri and Volín, 2010; Volín et al, 2013; Skarnitzl and Šturm, 2014). All the studies mentioned so far focused either on speech perception or on the presence of general phonetic/phonological characteristics of Czech (e.g. stress placement or spectral quality) in the Czech realization of English. The dependence of the possible Czech-to-English interference on the more highly levelled linguistic characteristics

(i.e. Czech lexicon) has been studies only in part by Růžek (2013). Generally, it can be stated that the parameters of Czechs' perception of English and the exact features of its realization by Czech speakers have not been fully explored and – with regard to this fact and to the supposition that English is likely to be even more widespread among native speakers of Czech – further research is desirable in order to deepen the existing knowledge of these phenomena.

3 MATERIAL AND METHOD

3.1 The hypotheses of the current study

Based on what was stated in the theoretical background, we can expect influence of the Czech lexicon on the realization of English words by Czech speakers. This expectation is based on two aspects:

a) the possibility of extension of Flege's Speech Learning Model (1995) on suprasegmental level
b) the differences in inherent temporal structures of Czech and English

Apart from the effect of analogues, we also expect that duration will play its role as one of the cues of prominence and that the Czech rules for stress placement will influence the temporal structures in the non-native realizations.

We thus propose the following hypotheses: Czech speakers will imitate the temporal structures of English words accurately when these words do not have analogues in the Czech lexicon while English words with analogues in the Czech lexicon will have temporal structure more similar to Czech. Czech speakers will also assimilate the temporal structures of English words to the Czech rules for stress placement.

3.2 Material

The research part of the current thesis focused on the segmental duration of specific English words in the pronunciation of Czech speakers. Since the emphasis was placed on the effect of the Czech lexicon on pronunciation, the main criterion considered in the selection of appropriate words was the presence or the absence of their analogous words in the Czech language. The list of the words that were used in this research consisted of two basic groups – words with an analogue in the Czech language (e.g. *abstract* /æbˈstrækt/) and words without a Czech analogue (e.g. *abjure* /æbˈdʒʊə/).

Other important parameters taken into account were the number of syllables, the placement of word stress and the quality of the vowel in the unstressed syllable (either full pronunciation or reduction to schwa). Because of the research extent affected by other variables and by the amount of analysed phonemes, only two-syllable words were used whereas words of three and more syllables were left out.

Three groups of words were established according to the presence of a specific vowel in one of the syllables. These vowels of interest were the open front unrounded vowel /æ/, in the spelling represented by the letter "a", the mid front unrounded vowel /e/ represented by the letter "e", and the open-mid back rounded vowel /ɒ/ represented by the letter "o". The target words for the analysis were chosen in such a manner that each of these vowels appeared once in the first syllable and once in the second syllable (3 vowels × 2 positions). Each of these vowels was also once under stress and once in an unstressed positions (thus 3 vowels × 2 positions × 2 stress levels). Additionally, when the vowel was in the unstressed position, words with both full (/æ e ɒ/) and reduced pronunciation (the mid central vowel, /ə/) were included, as in the word *complain* (/kəmˈpleɪn/ × /kɒmˈpleɪn/) or in the pair *permit* vs. *penult* (/pəˈmɪt/ × /peˈnʌlt/). In total, 18 different structural types were used.

Further, each of these 18 types included two structurally similar words, one member of each pair having an analogue in Czech and the second without it. The structural similarity consisted mainly in the identity (or in few cases phonemic similarity) of the syllable containing the variable vowel. Whenever possible, it was also intended to use pairs of words where the other syllable comprised vowels of the same vowel length (e.g. the presence of diphthongs in the first syllables of *Roman* /ˈrəʊ mən/ and *layman* /ˈleɪ mən/).

The absolute majority of the analysed words with Czech analogues were of Latin origin and with the exceptions of *peruke* and *complot* they are standard items in the English lexicon. All the Czech analogues are also commonly used. They appear in the same form as their English counterparts (adjusted to the Czech spelling and pronunciation rules and with respective suffixes) and they bear the same or similar meaning as their English counterparts (e.g. abstract - *abstrakt(ní)*, perfect - *perfektní*).

The words that do not have any analogue in Czech are usually also of Latin origin, but they have not been integrated into the Czech lexicon. When there are English words in the list that have the same etymological source as their Czech translations, but the phonological forms of these words changed substantially so that they do not resemble each other, we considered them as words without Czech analogues. It was the case of the word *abbey* and its Czech translation *opatství*, which have shared origin in the Latin word *abbatia*, but their respective forms diverted significantly. In case of

the word *epact* there is a Czech translation of a similar form (*epakta*); the Czech word is, however, an astronomical term and is infrequent in the Czech lexicon and therefore *epact* was also considered a word without an analogue.

In the course of the selection of the English words, it was also necessary to avoid words that do not have an analogue in the Czech language, but whose form is similar to a different English word which does have a Czech analogue. Words such as *seaman* and *Ascot*, which were originally included in the list, had to be omitted due to their resemblance to the words *semen* and *escort*, and they were substituted by *layman* and *ergot*.

The final number of words chosen for the analysis was 36; in the word *abstract* (/æbˈstrækt/), the vowel /æ/ was taken into account in both first syllable unstressed and second syllable stressed position. It could thus be analysed as two structural types. On the other hand, we added the word *absent* and also the word *torment* was recorded in both possible options (/ˈtɔː ment/ and /tɔːˈment/) and the realization with final stress duplicated the word *foment* in our scheme. The pronunciation of all the words follows the *Longman Pronunciation Dictionary* (LPD; John C. Wells, 2008) with the exception of *complot*, whose pronunciation is derived from the *Oxford English Dictionary* (Oxford English Dictionary, 2015). In case of some unreduced vowels in unstressed positions, the less common pronunciation provided by LPD had to be chosen (e.g. /ˈɜːg ɒt/ instead of more frequent /ˈɜːg ət/). The full list of words along with their pronunciations is provided in the Appendix.

3.3 Procedure

The whole procedure of recording took place in January 2015 in the recording studio of the Institute of Phonetics at the Faculty of Arts, Charles University. The recordings were obtained at a 32-kHz sampling rate and 16-bit quantization. All the words were recorded by a female British speaker with the pronunciation elicited to serve the purpose of the research (this relates especially to the multiple pronunciation variants of some words). The native speaker's recordings were considered the referential standard of English pronunciation, to which the pronunciation of Czech speakers was later compared.

In total, there were 16 target speakers analysed in the research, 7 of which were female and 9 male. All the participants were young native speakers of Czech (aged between 20 and 30 years), mostly university students and none of them was a university student of English, another language or general linguistics.

Prior to the recording itself, the subjects were given 8 test samples so that they could familiarize themselves with the form of the recording and the volume settings. The recording session of each speaker was divided into two blocks comprising the same number of items. The participants were given a three minute break between these blocks in order to avoid the effect of fatigue on their pronunciation. The second advantage of the break was the possibility to minimize the chance that a pair of identical words (but with a different pronunciation) would appear one after the other. The words which appeared twice in the list of items – i.e. *abstract, torment* and *complain* – were put in different groups; in the case of the word *contrast*, which was pronounced in three various ways, two versions were put in the first group and one in the second group. In addition, the distribution of the words guaranteed the same number of the words with analogues in both blocks.

The recordings of the native speaker were adjusted for the purpose of the research in the following way. A silence period was generated and placed after each word, so that the whole sequence containing one recorded word was 2.5 seconds long. These sequences were copied one after another into passages comprising five sequences (total duration of 12.5 seconds). The Czech speakers were told to listen to the items and repeat the word which they heard as precisely as possible into the silence provided after each of the five words. Short instrumental desensitization tracks were played after each passage in order to avoid perceptual interference of words from adjacent passages. After the recording session, the speakers were asked to fill in a brief questionnaire concerning their English language background (see the Appendix).

3.4 Analysis

At the end of the whole recording session, the recordings that were collected contained five samples of all 36 words provided by each speaker. Due to the amount of segments to be analysed, the last samples of each recorded word were omitted from the analysis. It can also be expected that the speakers might have anticipated the end of each passage,

which would probably affect the pronunciation of the last word. Therefore, the analysis was based on 2,304 words in total (16 speakers × 36 words × 4 repetitions). The recordings were automatically segmented into words and phones by the Prague Labeller algorithm (Pollák, Volín and Skarnitzl, 2007). All segment boundaries were manually checked in Praat (Boersma and Weenink, 2015) and corrected where necessary. The segmentation was performed according to the guidelines in Machač and Skarnitzl (2009). A Praat script was run on the TextGrid files with the corrected boundaries, which yielded a table, imported into Excel, with the complete data on the segmental duration.

After a final adjustment of the acquired data (see Results), an analysis of variance (ANOVA) was applied on the required variables. In the analysis, the dependent variable was a duration difference between the Czech and the native realizations of words, syllables, specific sequences and selected segments; the independent variables were Czech analogues (their presence × their absence), type of vowel realization (stressed × unstressed full × unstressed reduced), stress placement (initial × final), repetition (1 × 2 × 3 × 4) and the individual speakers (F1-F7 and M1-M9). Figures and tables of selected combinations of the variables were acquired from the software *Statistica* (Statsoft Inc., 2014) and Tukey's post hoc test was applied on the acquired data to establish statistical significance of the differences.

4 RESEARCH PART

4.1 Results

The total amount of observed words was 2,304. Nevertheless, it was necessary to eliminate those recordings of imitated words which differed completely from its native form and whose realization would negatively affect their temporal structures. The elimination was based on the following criteria:

We eliminated words which:

- did not include final plosive (mostly /t/), which represented a substantial part of the duration of the word
- did not include other segments of significant duration
- had their segments in inaccurate order (e.g. /fɒks trɒt/ → /fɒks tɒrt/)
- had one consonant replaced by another consonant with a different manner of articulation

We did not eliminate words which:

- included /r/ when the rhotic variety allows such realization (e.g. /tɔːrment/)
- differed in the spectral quality but not in the quantity of their vowel
- had a consonant replaced by another consonant with the same manner of articulation (e.g. /kɒm ent/ → /pɒm ent/)
- did not include schwa when there is a possible variety with a syllabic consonant (e.g. /ped l̩/)
- did not include schwa as reduced variations in unstressed syllables (e.g. /km 'pleɪn/)
- did not include /r/ after a plosive which is spirantised, thus temporally compensating for this omission (e.g. /kɒn taːst/)

The number of excluded items was 34 in total. In order to maintain objectivity of the results, we also excluded one speaker from the final data. The mean duration of speaker M8 was 91 ms greater than the duration of the native realization while the second longest average duration was only 32 ms greater (F7). This difference was probably caused by intentional lengthening of the final consonants (mainly alveolar plosives), which could be perceived from the recordings of speaker M8. Results of no

other speaker revealed any excessive deviation from the rest. Therefore, our analysis includes the realizations of 36 target words by 15 speakers, with the final amount of analysed words being 2,126 (15 speakers × 36 words × 4 repetitions minus 34 discarded tokens).

In the analysis of the results, we approached the acquired data from three perspectives. Firstly, in accordance with our main hypothesis we observed the temporal structures of the English words in dependence on presence (A+ words) or absence (A– words) of a Czech analogue. Secondly, we focused on the duration of vowels /æ/, /ɒ/, and /e/ and their reduced realizations in identical consonantal environments. The vowels and sequences *AB, MENT* and *CON/M* were also approached with regard to analogues. Thirdly, we analysed the durations of syllables in order to determine a possible effect of Czech stress placement on the non-native realizations. Finally, we tried to find out whether some speakers differed significantly from others in their ability to imitate temporal structures of the native speaker.

The first major result of the research part of the current thesis is that the average word duration was shorter in the Czech realizations by 39 ms (see Table 1). 25 out of the 36 words were shorter than their L1 English model counterparts. It can be seen from the table that the most lengthened word *Roman* was longer by 44 ms (8% of the native speaker's duration), on the opposite extreme, *penult* was shortened by 139 ms (17% of the native speaker's duration). While the smaller duration of the first syllable was detected in 20 words, this amount was 28 in case of the second syllable.

Word	**Czech analogue (yes/no)**	**Word duration difference (average)**
Roman	Yes	44.3
foxtrot	Yes	43.4
layman	No	37.2
moment	Yes	32.8
abbey	No	28.0
ballot	No	23.5
complain (reduced)	No	21.2
pedal	Yes	14.6
abjure	No	3.1

absurd	Yes	2.6
pebble	No	0.2
abhor	No	-2.6
abstract (stress 1)	Yes	-3.4
perfect	Yes	-10.1
detract	No	-30.5
absent	Yes	-30.9
complain (full)	No	-32.3
payment	No	-32.9
torment (stress 2)	No	-33.3
contrast (stress 1)	Yes	-45.9
comment	Yes	-53.7
faggot	Yes	-63.3
contrast (stress 2; full)	Yes	-68.0
contrite	No	-68.7
foment	No	-79.2
peruke	Yes	-83.2
ergot	No	-84.9
complot	Yes	-87.9
torment (stress 1)	No	-89.9
cement	Yes	-92.4
permit	No	-100.9
contrast (stress 2; reduced)	Yes	-103.9
forgot	No	-104.3
abstract (stress 2)	Yes	-107.5
epact	No	-116.6
penult	No	-139.2
All words		**-39.3**

Table 1: Mean duration difference of individual words (in ms) and their A+/A– classification.

4.1.1 The effect of analogues on the temporal structures

We will now focus closely on the variable of Czech analogues. The word duration tends to be almost identically shortened in words with and without Czech analogues (A+: –36 ms; A–: –42 ms; F(1, 2124)=1.71; $p > 0.05$). Similar results were shown for the duration of the second syllable – shortening by 36 ms in A+ words and by 39 ms in A– words ($p > 0.05$). The duration of the first syllable was imitated more accurately on the whole and with a statistically significant difference between the two conditions – it was on average 3 ms greater in A+ words and 6 ms smaller in A– words (F(1, 2124)=23.46; $p < 0.001$).

The analysis of the influence of A+ words and A– words in relation to the stress placement provided us with the following average results.

The mean duration of words with initial stress as seen in Figure 1 was 8 ms smaller than in the native realization when there was an analogue in Czech, while it was even smaller – with a mean difference of 33 ms – when no analogue was present in Czech (F(1, 2122)=39.45; $p < 0.001$). As regards the words with final stress, the shortening was on average 68 ms (with Czech analogues) and 48 ms (without Czech analogues) ($p < 0.001$). The differences are highly significant.

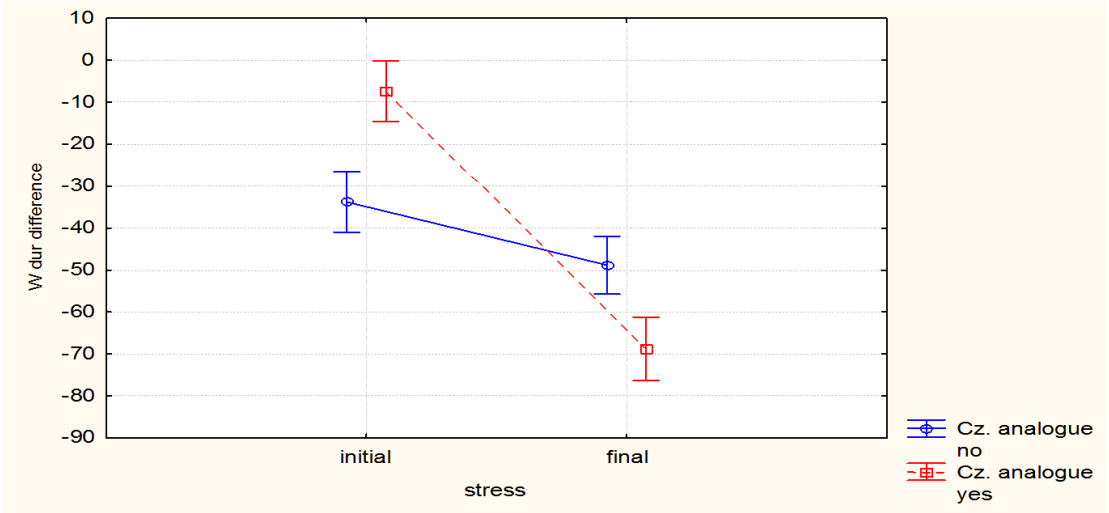

Figure 1: Mean word duration difference between L1 and L2 productions (in ms) with regard to analogues and stress placement.

If we focus on the first syllable only, its mean duration in words which have initial stress and also have an analogue was 6 ms greater than in the native realization. When the stress was in the initial position but there were no analogues, the first syllables were of the same mean duration as in the native realization ($F(1, 2122)=1.28$; $p > 0.05$). In the case of final stress, the mean duration was native-like in words with analogues and smaller by 12 ms in words without analogues ($p < 0.001$).

As far as the duration of the second syllable is considered, the results revealed that – when the stress was on the first syllable – A+ words were 13 ms shorter and A– words were 33 ms shorter than in the native realization ($F(1, 2122)=59.0$; $p < 0.001$). In words with the final stress, the mean duration of second syllable was smaller by 68 ms in words with analogues and by 37 ms in words without analogues ($p < 0.001$). In both relations, the differences are highly significant.

We also analysed the effect of the A+ and A– words on duration with regard to the repetition of the word. In Figure 2, it can be seen that both types of words underwent a slight amelioration: in A+ words by ca. 5 ms, and in A– words by ca. 10 ms in the first three repetitions followed by a slight deterioration. Similar results were acquired for the duration of the first syllable where the repetition did not influence words with analogues and slowly ameliorated words without analogues (by ca. 5 ms). Only in the duration of second syllables the A+ words were closer to the native duration than the A– words. Words without analogues ameliorated a little more than words with analogues, but only in the first three repetitions. None of the differences concerning repetition was found statistically significant ($F(3, 2118)=0.17$; $p > 0.05$).

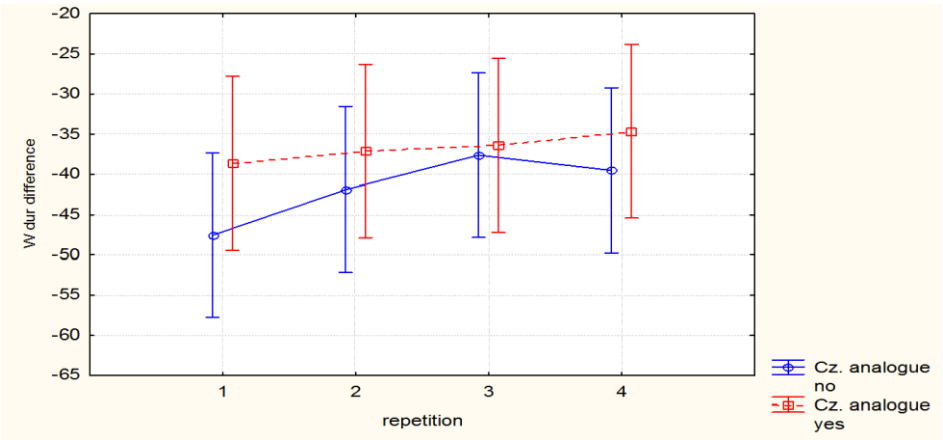

Figure 2: Mean word duration difference between L1 and L2 productions (in ms) with regard to analogues and repetition.

4.1.2 The temporal structures of selected sequences and their vowels

In addition to the above results based on the data from all the words together, temporal structures of three selected vowels were analysed more closely in the following sequences:

- /æ/ in sequence *AB* (as in *abstract*)
- /e/ in sequence *MENT* (as in *torment*)
- /ɒ/ in sequence *CON/M* (as in *contrast* or *complain*)

Each vowel was observed in three different realizations – in full realization as a part of a stressed syllable, in full realization as a part of an unstressed syllable, and in reduced realization as a part of an unstressed syllable. The A+ words and A– words dichotomy was also taken into account. Apart from the vowel duration, we analysed also the duration of the whole sequences.

Results for AB sequence

The vocalic duration in the sequence *AB* did not reveal any statistically significant differences with regard to analogues ($p > 0.05$). Therefore, the analogue classification is not discussed in closer detail. However, differences appeared with respect to the main effect of vowel type (stressed, unstressed full, unstressed reduced); $F(2, 351)=68.50$; $p < 0.001$. In Figure 3a, we can see that the average duration of the vowel /æ/ in stressed syllables was greater by ca. 8 ms than the duration of its native model. In unstressed syllables with full realization of the vowel, the duration was smaller by 21 ms, and the reduced vowels were on average 3 ms smaller. All these prominence-related differences proved to be statistically highly significant ($p < 0.001$).

As far as the duration of the whole *AB* sequence is considered (see Figure 3b), we observed a significant interaction of the prominence level with the presence or absence of analogues ($F(2, 351)=74.67$; $p < 0.001$). The sequence duration in stressed syllables of A+ words was native-like while in A– words it was greater by 31 ms ($p < 0.001$). When the sequence was unstressed and contained a full vowel, its duration was slightly greater in A+ words (by 12 ms) and accurate in A– words ($p < 0.05$). The sequences with reduced vowels were only slightly shorter in words with analogues but longer by 62 ms in words without analogues ($p < 0.001$).

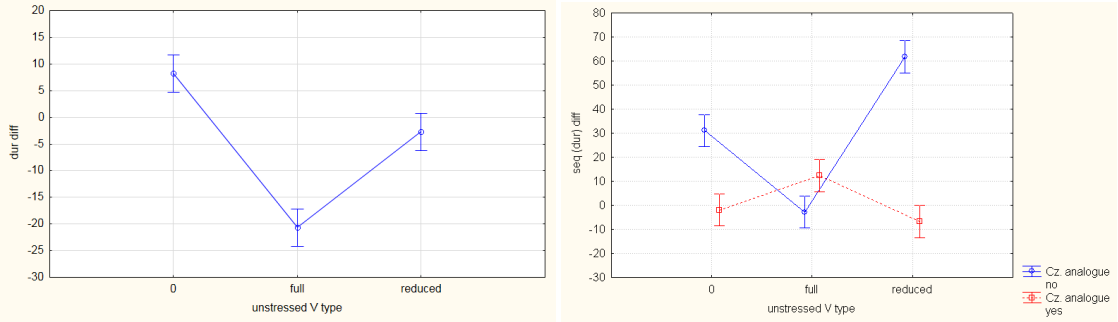

Figure 3a: Mean segment duration difference between L1 and L2 productions of vowel /æ/ in *AB* sequences with regard to vowel type.

Figure 3b: Mean sequence duration difference between L1 and L2 productions of *AB* sequence with regard vowel type and analogues.

Results for MENT sequence

In Figure 4a, the results for the mean duration of vowel /e/ in the sequence *MENT* are shown. On the whole, the duration was larger in the imitations than in the native realization. In stressed position, the duration of the vowel was similarly greater in both variations (in A+ words by 24 ms; in A– words by 22 ms) when compared to the native realization ($p > 0.05$). The mean vowel duration was greater also in words with full realization in unstressed position – by 24 ms in words with analogues and by 10 ms in words without analogues ($p > 0.05$). When reduced, the vowel was longer by 23 ms in A+ words and by 26 ms in A– words ($p > 0.05$). None of the differences between words with and without analogues was statistically significant, and there was no significant interaction with the prominence level.

Duration of the whole sequence *MENT* is presented in Figure 4b. There was a significant interaction, as the stressed position was the only one in which A+ and A– words were distinguished: $F(2, 353)=12.35$; $p < 0.001$. When the sequence was stressed, its duration was smaller by 84 ms in A+ words and by 48 ms in A– words ($p < 0.05$). In an unstressed position with a full vowel, the sequence was shorter by 64 ms in words with analogues and shorter by 92 ms in words without analogues ($p > 0.05$). The sequence *MENT* with reduced vowel was on average longer by 15 ms when having an analogous word but shorter by 14 ms when without an analogous word ($p > 0.05$).

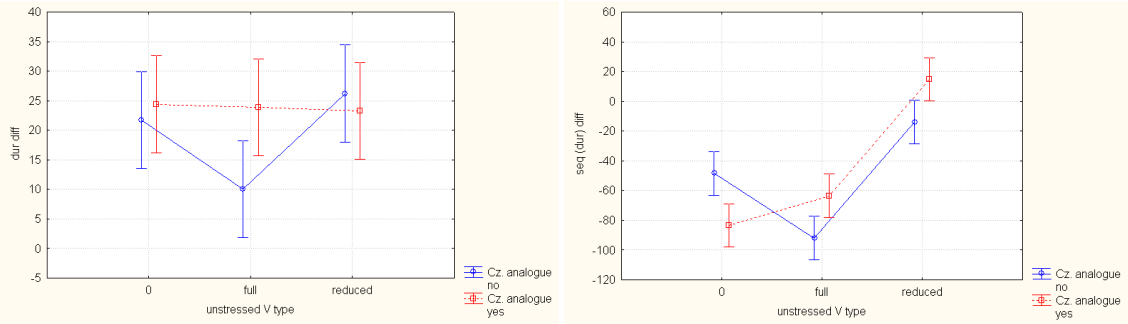

Figure 4a: Mean segment duration difference between L1 and L2 productions of vowel /e/ in *MENT* sequences with regard to vowel type and analogues.

Figure 4b: Mean sequence duration difference between L1 and L2 productions of *MENT* sequence with regard vowel type and analogues.

Results for CON/M sequence

The results for the vowel duration in the sequence *CON/M* are shown in Figure 5a. The interaction between the main effects was significant: F(2, 340)=28.19; $p < 0.001$. The mean duration of the vowel /ɒ/ in stressed position was greater in both A+ words (by 6 ms) and A– words (by 17 ms) ($p > 0.05$). With full realization in an unstressed position, the vowel was slightly shorter than in the native realization in both variations – by 2 ms when there was an analogue in Czech and by 5 ms when there was no analogue ($p > 0.05$). However, reduced vowels in sequence *CON/M* were also almost accurate in words without analogues (shorter by 3 ms) but longer by 28 ms in words with analogues ($p < 0.001$).

We can see in Figure 5b that – when the duration of the whole sequence was taken into account – the *CON/M* sequence in stressed position was shorter by 18 ms in A+ words and by 47 ms in A– words ($p < 0.001$), which explains the significance of the interaction: F(2, 340)=16.39; $p < 0.001$. When the sequence was in an unstressed position and contained a full vowel its duration was similarly shorter in both conditions – by 17 ms in words with analogues and by 15 ms in words without analogues ($p > 0.05$). When there was an analogue in Czech, the sequence *CON/M* with a reduced vowel was on average longer by 13 ms, whereas it was longer by 29 ms when the analogue was absent ($p > 0.05$). The reduced position thus seems to be associated with greatest lengthening compared to the other positions.

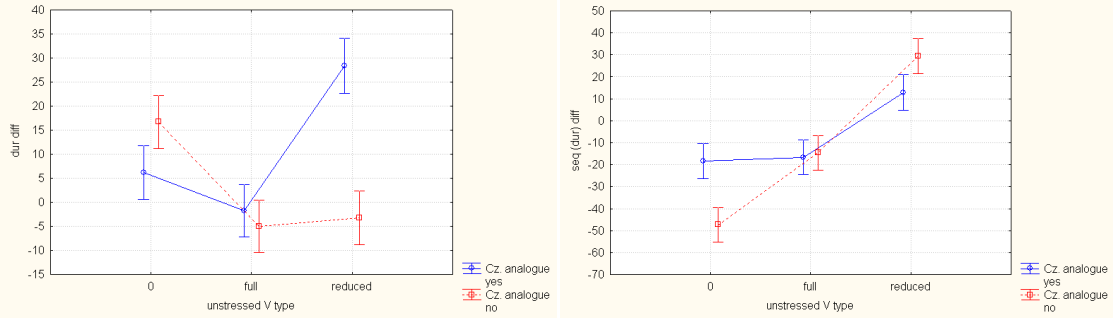

Figure 5a: Mean segment duration difference between L1 and L2 productions of vowel /ɒ/ in *CON/M* sequences with regard to vowel type and analogues.

Figure 5b: Mean sequence duration difference between L1 and L2 productions of *CON/M* sequence with regard vowel type and analogues.

4.1.3 The differences in the results of individual speakers

The analysis of the ability of Czech speakers to imitate temporal structures of English words and syllables also revealed differences among individual speakers. In Figure 6a we can observe that the majority of the speakers reproduced English words with a smaller duration than the native speaker and that only speaker F7 had the mean duration of both A+ and A– words greater by more than 20 ms. Most accurate were speakers F1, F2 and M4. It is apparent that, with the exception of speakers F4 and M4, the female speakers (F1-F7) were a lot more accurate than the male speakers (M1-M7, M9). The mean diversion of the female speakers was 5 ms while the male speakers shortened their realizations by 70 ms on average.

While the imitation of the first syllable with regard to duration was more accurate (see Figure 6b), i.e. closer to zero and with only three speakers diverting by more than 20 ms on both sides, 10 out of the 15 speakers diverted by more than 20 ms when imitating the duration of the second syllable (see Figure 6c). In the measured durations (word, first syllable, second syllable) it was speaker F2 who appeared to be the most accurate in the imitation.

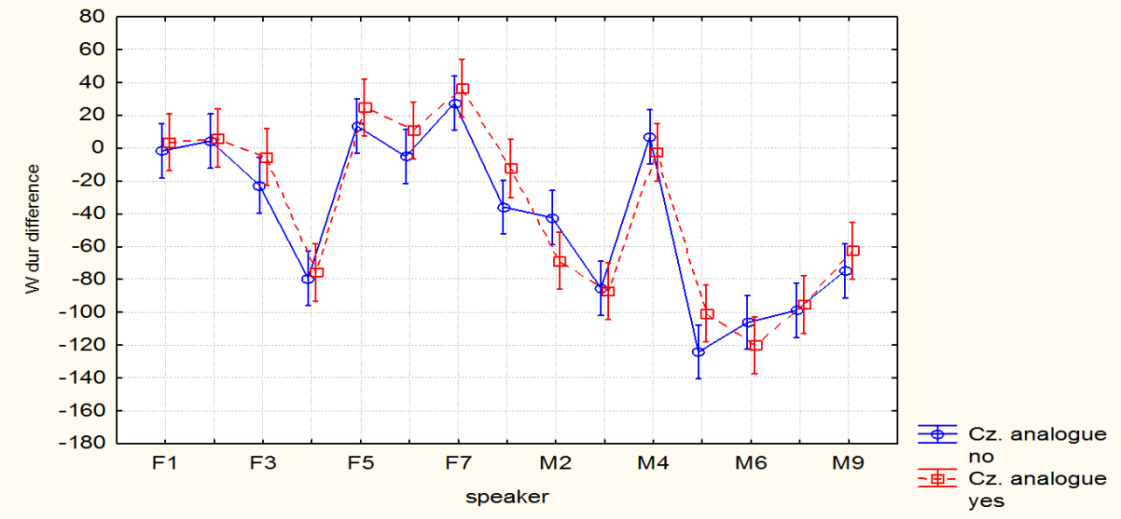

Figure 6a: Mean word duration difference between L1 and L2 productions with regard to analogues and individual speakers.

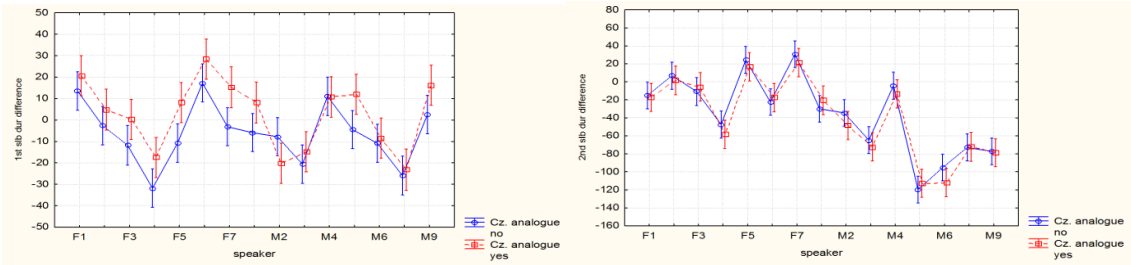

Figure 6b: Mean first syllable duration difference between L1 and L2 productions with regard to analogues and individual speakers.

Figure 6c: Mean second syllable duration difference between L1 and L2 productions with regard to analogues and individual speakers.

4.2 Discussion

As was already stated in the previous subchapter, one of the observations acquired from the research is that the temporal structures of English words were shorter in realizations of Czech speakers. It is difficult to determine the reasons for this phenomenon, but we can assume that the speakers may have subconsciously or deliberately sped up their realizations of the words in order to fit them into the set silence periods. As the shortening appeared to be a general matter, we accepted this finding as a baseline for the other acquired data. Mean word duration difference of Czech realization was set according to the results: –39 ms in comparison

with the native realization. Ideally, each syllable should contribute equally to the shortening, which means that they should be shorter by almost 20 ms. Words would thus be diminished, yet structurally exact copies of the native model.

It can be stated that our hypothesis that Czech speakers will imitate the temporal structures of English words accurately in words without analogues and inaccurately in words with analogues was only partly confirmed. The effect of analogues was found in the first syllable where the Czech speakers indeed diverted more from the expected realization (i.e. equal shortening of both syllables) when they knew the words from the Czech lexicon. Moreover, the difference between the conditions was shown to be statistically significant. The fact that the first syllable was affected by the variable of analogues could be caused by the tendency of Czech speakers to move lexical stress temporal prominence to the first syllable when the second is stressed. The stress on the initial syllable is obligatory in Czech and the speakers lengthened it more in the words that actually appear in their native language. No such divergence was found in the second syllable.

The results of the analogues effect in interaction with stress placement are definitely of interest. In words with final stress, there were highly significant differences between all the realizations of A+ words and A– words. While the speakers lengthened the first syllables and shortened the second syllables of the words which also appear in Czech, they imitated temporal structure more accurately when the words should be unknown to them (or, more precisely, when there should be no interference of Czech phonological structure of the analogous words). Interestingly enough, the relative lengthening of the first syllable was present also in words with initial stress, i.e. in words where no cross-syllabic transfer of prominence is needed. In these stress-initial words, we also observed a stronger resistance against the stress shift in words without analogues. Even though the Czech speakers found it difficult to keep the temporal prominence of the native model, they seem to have managed to do so better in words that do not appear in the Czech lexicon. Therefore, we can assert that in the imitation task the presence of analogues had a deteriorating effect on the way Czech speakers distributed syllable duration in dependence on the stress placement.

Since all the analysed words contained only two syllables the observations of syllabic prominence and its possible shift provided us with only two options. Future studies can

show how the presence of analogues would affect duration and its possible contribution to prominence shift in words with more than two syllables. Apart from stressed and unstressed syllables, a category of secondary word stress could be added, providing for a better insight into the issue of Czech and English interference.

Some tendencies were observed also in relation to repetition. Words in their first realization are more diverted from the native model when they do not have an analogue. While all words are ameliorating gradually, the progress is more significant in words without analogues. In other words, it seems that Czech speakers have to get used to the temporal structures of unknown words, which might be caused by a false recognition of the word when heard for the first time. Their realization, however, is then shown to improve better than in the case of the A+ words (in the first three repetitions). The temporal structures of words known from Czech seem to be less flexible. Nevertheless, we can speak only about tendencies.

According to the findings about the effect of stress placement on the duration, lengthening was expected in the sequences *AB* and *CON/M* as they were identical with the first syllable. On the other hand, the sequence *MENT*, which was in the final position, was expected to be shorter. Similar tendencies were looked for also in terms of the individual vowels and were supposed to be more apparent in words with analogues. However, the results of the selected sequences and their vowels do not suggest similar findings as the analysis of all words, i.e. the resistance against interference of Czech and English in words without analogues. The duration of the vowels often seemed to be unaffected by the variable of analogues and when it was affected, it was the duration of A– words that differed more from the expectations. Moreover, these deviations did not systematically appear with respect to the observed vowel types – they were found in the stressed and both unstressed variations.

The results are similarly unstable for the duration of the whole sequence. With regard to the analogues Czech speakers either imitate their durations equally or against our expectations. Only in the realization of *CON/M* under stress and in the realization of *MENT* with reduced vowel, the sequences kept their presupposed mean duration difference of –20 ms. In the case of *MENT*, however, this might have been caused by the different perception of syllabic boundary in Czech and English, a problem which is also mentioned in the discussion of Růžek's study (2013). English tends to classify

intervocalic consonant as a coda of the first syllable. Therefore, the words *moment* and *payment* have their syllabic boundary between *m* and *e* and the segment /m/ is a part of the first syllable. On the other hand, Czech as a language containing more open syllables (i.e. syllables of CV structure) and thus the boundary in the Czech analogue of *moment* places the segment /m/ at the beginning of the second syllable. As the duration of a segment in syllabic coda tends to be smaller than the duration of the same segment at the beginning of a syllable, the durations of the segment /m/ in the sequence *MENT* are expected to differ. If a word with the segment /m/ as a part of the second syllable would be selected instead of *moment* and *payment*, it is likely that the duration differences between the Czech and the native realizations would be even more than the 20 ms. Generally, our analysis thus does not bring any conclusive findings about the effect of the analogues on the duration of the selected sequences and their vowels.

It is, however, apparent from the results that the vowel duration change does not seem to have a substantial effect on the change of the sequence duration. When the vowels were unaffected by the analogue dichotomy, the sequence durations often differed substantially in this respect, and vice versa. For instance, the speakers had problems with imitating reduced vowels in sequence *MENT*. They realized these vowels with greater duration than the native speaker. However, this was not supported by the duration of the whole sequence, which was very accurate. These findings suggest that duration of not only vowels but also consonants may differ in the Czech realization and can be dependent on the analogues. It might be the aim of further research to find out the relevance of this tendency.

If the variable of analogues is not considered, interesting results are revealed with respect to stress placement, which was our second hypothesis. Volín and Weingartová (2014: 181) observed that Czech speakers "deviate more substantially from the canonical placement of stress, frequently shifting it to the first syllable of a word." In their study, the acoustic correlates attributing to stress marking were mainly pitch and intensity, while the temporal structures were quite accurate. The results of the current study show a strong tendency to a prominence shift also in the domain of time.

It has been already stated that, if the general shortening is taken into account, both syllables should be ideally shortened by an equal amount. However, the first syllable is realized accurately (i.e., instead of the ca. 0 difference we may treat is as lengthening if we take into account the –20 ms baseline) while the second syllable is on average significantly shortened (even if related to the baseline). This is most probably the effect of the obligatory initial stress placement in Czech.

Apart from this shift we can also see that there is a highly significant difference between the temporal distribution in words with initial and final stress. When normalized to the baseline, words with initial stress are substantially longer than words with stress on the second syllable and similar results are revealed also for the durations of individual syllables (first or second). Generally expressed, Czech speakers exaggerated the durational prominence of the first syllable in words that already had initial stress and they maintained the duration of their second syllable. This resulted in the overall lengthening of these words. On the other hand, the first syllable was slightly lengthened (possibly due to stress shift) in words with final stress while the second syllable (losing its prominence) was extremely shortened. These changes caused that the mean duration of these words was smaller.

It is of interest that the major focus of the change was in both cases on the stressed syllable. The initial stressed syllable was not perceived as sufficiently long and analogically the speakers seemed to have the urge to shorten the second syllable when it was stressed. The perceptual assessment of the recordings did not seem to show similar tendencies in terms of other acoustic cues of syllabic prominence. It is thus another possible task for further research to examine the potential differences in pitch, intensity and spectral quality.

We have also observed that female speakers were in general more accurate in imitating temporal structures of words. The results, however, do not much coincide with the expectations based on the speakers' background of English. The questionnaires, included in the Appendix, revealed that only two speakers started to learn English after the age of 10. Surprisingly, one of them was the speaker F2 whose imitation was not only the most accurate but also most consistent with respect to analogues and positions of stress. On the other hand, the only speaker who had spent a significant time

in an English speaking country had the mean word duration almost 100 ms shorter than the native speaker.

The speakers did not differ significantly in other respects. All were university students and none of them was in a daily contact with native speakers. Examining a broader group of speakers in terms of their level of experience with English could also show more representative results and reveal more information about the interference of Czech and English.

In the course of the research, we have come across several difficulties. Because of the multiple requirements on the word characteristics (with or without Czech analogues, the vowels involved, word stress, structural similarities), the selection of units for analysis was not an easy task and might be disputable in some cases. The words *permit* and *penult,* for instance, do not exist in Czech lexicon as such but are present as the bases for the words *permitivita* and *penultima. Penult* may have caused also another difficulty because it resembles the Czech word *penalta*. Similar problem was also present with the word *ballot* and its false Czech counterpart *balet.* An eschewal of these ambiguities in the word selection in future research might contribute to more accurate results.

Comparing the realizations of similar words with short and long vowels in their Czech analogues could be another contribution to the studies of Czech English. Only 2 out of the 17 words with analogues in the list contained an inherently long Czech vowel (*Roman – román* and *pedal – pedál)*. Moreover, the long vowels form the Czech words were in both cases in the reduced position of their English counterparts.

In conclusion, our hypothesis that Czech speakers will imitate the temporal structures of English words accurately when the words do not have analogues and with more marked temporal diversions in words having analogues was confirmed only partly. However, we found significant differences in the ways Czech speakers distributed syllable duration in dependence on the stress placement and these differences are consistent with the direction of the hypothesis. It was also shown that Czech speakers generally tend to assimilate temporal distribution in the words with final stress to the Czech stress placement, i.e. to shift the durational prominence to the first syllable. Finally, we observed that female speakers tend to be more accurate in their imitations of temporal structures then male speakers.

5 CONCLUSION

The aim of the current thesis was to determine whether there was an effect of the presence and absence of Czech analogues on the temporal structures of English words in the realization of Czech speakers. Our hypothesis was based on the extension of Flege's Speech Learning Model (SLM; 1995) to the suprasegmental level of speech and the differences between the temporal structures in Czech and in English. In this respect, the current thesis follows the diploma thesis of Jan Růžek (2013), who focused on phone durations in multi-syllable words with Czech analogues in Czech English. Apart from the main hypothesis, we were also interested in the possible effect of the Czech stress placement on the way Czech speakers imitate the temporal structures of English words.

In the Theoretical background, the main areas of our focus were described. The first subchapter, which dealt with foreign accent, presented the basic definition of the term. The matter of foreign accent was approached from the perspectives of segmental and suprasegmental level. The major emphasis was placed on the description of SLM and its possible extension to the suprasegmental level, which had been confirmed by studies of McAllister et al (2002) and Mennen (1999; cited in Piske, 2001). We also focused on the factors that affect foreign accent emergence and perception. These factors were observed on the side of the speakers, on the side of the listeners and also on the side of the research procedure.

Two selected phenomena were observed in detail in the following subchapter. The first of them was lexical stress. We introduced the four major acoustic cues of lexical stress – pitch, intensity, duration, and spectral quality and summarized the findings about the strength of the individual cues. The second phenomenon of our focus in this subchapter was duration. Apart from its function as one of the correlates of lexical stress, duration plays a role also on other levels of speech (e.g. semantic, syntactic, physiological). Because of the purpose of the current thesis we focused on the inherent phonological duration of segments and the duration of vowels in dependence on the voicing classification of the following consonant. As lexical stress and duration are areas which differ significantly in Czech and in English, a significant effect of interference of the languages was expected.

The final part of the Theoretical background dealt with the English of Czech speakers. The major emphasis was placed on the studies analysing Czech English with respect to lexical stress and duration as one of its acoustic correlates. The existing findings in this area were presented in this subchapter along with predications about the research part of the current thesis.

The first chapter of the empirical part, Material and Method, presented our hypotheses and described the process of the research. We selected 36 two-syllable English words, which were then recorded by a British native speaker. The native recordings were played to 16 Czech speakers, who were asked to imitate the words as precisely as possible. All the acquired recordings were automatically segmented into words and phones and the data for the Czech speakers were compared to those for the native realization.

In the subchapter Results, the temporal structures of the native speaker and the Czech speakers were compared. The analysis was made with respect to the analogue classification of the words as well as to the units that were compared. These units were whole words, their syllables, selected sequences and their vowels. The results showed a general tendency of Czech speakers to shorten their realization of English words. The analysis also deals with the differences among the durations of individual speakers' realizations. All the differences in durations were presented with regard to their statistical significance.

In Discussion, the results of the research were approached with regard to our hypothesis. It was shown that the hypothesis that the presence of analogues will affect the temporal structure of English words in realization of Czech speakers was partly confirmed. The word durations and syllable durations differences were indeed greater when the imitated word had an analogue. The results also confirmed the expected prominence shift from the second syllable to the first syllable as a consequence of the stress placement differences between English and Czech. Finally, the discussion presents suggestions for future research which may help to deepen the existing knowledge about the English of Czech speakers.

REFERENCES

Best, C. T. (1995). A direct realist view of cross-language speech perception. In: Strange, W. (ed) Speech Perception and Linguistic Experience: Issues in cross-language research. Timonium, MD: York Press.

Bissiri, M. P., and Volín, J. (2010). Prosodic structure as a predictor of glottal stops before word-initial vowels in Czech English. In: R. Vích (Ed.), 20th Czech-German Workshop - Speech Processing, Prague, pp. 23-28. Prague.

Boersma, P., and Weenink, D. (2014). Praat: Doing phonetics by computer (Version 5.4.08.). Downloaded on 14 December 2014 from http://www.praat.org/.

Bolinger, D. L. (1958). A theory of pitch accent in English. In: Word, 14, pp. 109-149.

Braun B., Lemhöfer, K., and Mani N. (2011). Perceiving unstressed vowels in foreign-accented English. In: Acoustical Society of America, 129.1, pp. 376–387.

Cruttenden, A. (2008). Gimson's Pronunciation of English. London: Hodder Education.

de Jong, K. (1995). Stress, lexical focus, and segmental focus in English: patterns of variation in vowel duration. In: Journal of Phonetics, 32 (2004), 493–516. Available on-line at www.sciencedirect.com.

Flege, J., and Hillenbrand, J. (1984). Limits on phonetic accuracy in foreign language speech production. In: Acoustic society of America, 76.3, pp. 708-721.

Flege, J. (1995). Second Language Speech Learning Theory, Findings, and Problems. In: Strange, W. (ed) Speech Perception and Linguistic Experience: Issues in cross-language research. 233-277. Timonium, MD: York Press.

Flege, J., Munro, M. J., MacKay I. R. A. (1995). Factors affecting strength of perceived foreign accent in a second language. In: Acoustic society of America, 97.5, pp. 3125-3134.

Flege, J., Freida, A., and Nozawa, T. (1997). Amount of native-language (L1) use affects the pronunciation of an L2. In: Journal of Phonetics, 25, pp. 169-186.

Frost, D. (2011). Stress and cues to relative prominence in English and French: A perceptual study. In: Journal of the International Phonetic Association, 41.1, pp 67-84.

Klatt, D. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. In: Acoustic Society of America, 59.5, pp. 1208-1221.

Kuhl, P. K., and Iverson, P. (1995). "Linguistic experience and the 'Perceptual Magnet Effect'". In: Speech Perception and Linguistic Experience: Theoretical and Methodological Issues. Ed. W. Strange. Timonium MD: York Press. Pp. 121-153.

Ladd, D. R. (2008). Intonational Phonology. Cambridge: Cambridge University Press.

Leather, J., and James, A. (1996). Second language speech. In: Handbook of Second Language Acquisition. San Diego: Academic Press. pp. 269-316.

Machač, P., and Skarnitzl, R. (2007). Temporal compensation in Czech?. In: Proceedings of the 16th ICPhS, pp. 537-540.

Machač, P., and Skarnitzl, R. (2009). Principles of Phonetic Segmentation. Praha: Nakladatelství Epocha.

Major, R. (2001). Foreign Accent: The Ontogeny and Phylogeny of Second Language Phonology. New York: Lawrence Erlbaum Associates.

McAllister, R., Flege, J., and Piske, T. (2002). The influence of the L1 on the acquisition of Swedicsh vowel quantity by native speakers of Spanish, English and Estonian. In: Journal of Phonetics, 30, 229-258.

Moyer, A. (2013). Foreign Accent: The Phenomenon of Non-native Speech. Cambridge: Cambridge University Press.

Piske, T., MacKay, I. R. A., and Flege, J. (2001). Factors affecting degree of foreign accent in an L2: a review. In: Journal of Phonetics, 29, pp. 191-215.

Plag, I., Kunter, G., and Schramm, M. (2011). Acoustic correlates of primary and secondary stress in North American English. In: Journal of Phonetics, 39, pp. 362-374.

Pollák, P., Volín, J. and Skarnitzl, R. (2007). HMM-Based Phonetic Segmentation in Praat Environment. Proceedings of the XIIth International Conference "Speech and computer – SPECOM 2007".

Růžek, J. (2013). Phone durations in multi-syllable words with Czech analogues in Czech English. Prague. Diploma thesis. Charles University in Prague, Faculty of Arts.

Skarnitzl, R. (2005). English Word Stress in the Perception of Czech Listeners. In: J.Čermák, A.Klégr, M.Malá & P.Šaldová (Eds.), Patterns, A Festschrift for Libuše Dušková. Praha: Kruh moderních filologů. pp. 183-194.

Skarnitzl, R. & Šturm, P. (2014): Assimilation of voicing in Czech speakers of English: The effect of the degree of accentedness. In: Research in Language, 12(2), pp. 199-208.

Southwood, H., and Flege, J. (1999). The validity and reliability of scaling foreign accent. In: Clinical Linguistics & Phonetics, 13, pp. 335-349.

Strange, W. (1999). Levels of abstraction in characterizing cross-language phonetic similarity. International Congress of Phonetics 1999, San Francisco, pp. 2513-2519.

Tajima, K., Port, R., and Dalby, J. (2007). Effects of temporal correction on intelligibility of foreign-accented English. In: Journal of Phonetics, 25, 1-24.

Volín, J., and Skarnitzl, R. (2010). Suprasegmental Acoustic Cues of Foreignness in Czech English. In: Proceedings of Odyssey: The Speaker and Language Recognition Workshop Odyssey 2010, pp. 271-278.

Volín, J., and Weingartová, L. (2014). Acoustic Correlates of Word Stress as a Cue to Accent Strength. In: Research in Language, 12(2), pp. 175-183.

Wells, J. C. (2008). Longman Pronunciation Dictionary (3rd ed.). Harlow: Longman.

Oxford English Dictionary (2015). Oxford English Dictionary: The definitive record of the English language. Accessed online at www.oed.com.ezproxy.is.cuni.cz on 8 January 2015.

StatSoft Inc. (2014). STATISTICA, version 10. www.statsoft.cz

RESUMÉ

Cílem bakalářské práce je zjistit, zda přítomnost český analogů má vliv na temporální strukturu anglických slov v realizaci českých mluvčích a zda čeština interferuje s angličtinou i z hlediska pozice slovního přízvuku. Práce je dělená na teoretickou a praktickou část. V teoretické části jsou představeny základní pojmy týkající se našeho zaměření, tj. především cizinecký přízvuk, slovní přízvuk a trvání. Jsou zde rovněž shrnuty dosavadní výzkumy, které se touto problematikou zabývaly. Praktická část se pak věnuje samotnému výzkumu, ve kterém jsou porovnávány temporální struktury anglických slov v realizaci britské rodilé mluvčí a 16 mluvčích, jejichž rodným jazykem je čeština. Výsledky výzkumu jsou pak předneseny a na závěr podrobeny diskuzi.

Na otázky, které si v práci klademe, lze nahlížet z více úhlů pohledu. Komplexně je tak pojímána i teoretická část práce obsažená v kapitole *Theoretical background*. Důležitými tématy jsou v ní cizinecký přízvuk, interference dvou jazyků na jejich různých rovinách, prominence jakožto projev slovního přízvuku, trvání z hlediska segmentálního i suprasegmentálního (neboli prozodického) a v neposlední řadě také čeština jako rodný jazyk mluvčích angličtiny.

Z hlediska cizineckého přízvuku se práce věnuje obecným znakům cizineckého přízvuku, ať už v rovině segmentální či prozodické. Velká část výzkumů týkajících se této problematiky se soustředila na její segmentální projevy, přestože se ukazuje, že cizinecký přízvuk je ovlivněn i prozodií. Jedním z autorů, kteří se do hloubky věnovali cizineckému přízvuku v angličtině, je americký lingvista Jim Flege. Právě o jeho Speech Learning Model (SLM; Flege, 1995) se práce nejvíce opírá. SLM dokazuje, že segmenty v druhém jazyce (L2), které jsou podobné nějakému segmentu v jazyce rodném (L1), se asimilují do podoby s těmito segmenty z L1 a jsou tak vysloveny nepřesně. Na druhou stranu pro segmenty, které jsou svou podobou vzdálené jakémukoli segmentu z rodného jazyka, si mluvčí vytvoří kategorii novou, realizovanou přesně. Přestože je SLM založen na zkoumání segmentální roviny, aplikovatelnost modelu i na vyšší rovinu prozodickou byla dokázána více výzkumy (McAllister, Flege & Piske, 2002 a Mennen, 1999 – citováno v Piske, 2001). Tato práce se snaží možnost tohoto rozšíření potvrdit s pomocí porovnávání slov s analogy a bez nich. Kapitola o cizineckém přízvuku se věnuje také faktorům, které ovlivňují realizační

a percepční projevy interference dvou jazyků. K těmto faktorům patří mimo jiné věk prvního kontaktu s druhým jazykem (AOL) nebo délka pobytu mluvčího v zemi, kde se tímto jazykem mluví.

Dalším velkým tématem teoretické části je rozbor dvou vybraných jevů, které hrají roli v cizineckém přízvuku. Prvním z těchto jevů je slovní přízvuk a s ním spojená prominence. Zatímco dříve byla za jediný projev slovního přízvuku považována pouze hlasitost, v současné době rozlišujeme čtyři projevy slovního přízvuku – výška tónu (základní frekvence F0), hlasitost (intenzita), trvání a kvalita vokálu. Síla jednotlivých projevů ve slovním přízvuku angličtiny byla zkoumána v mnoha studiích (např. Frost, 2011; Plag a kol, 2011). Vzhledem k tomu, že se čeština a angličtina v oblasti slovního přízvuku výrazně odlišují, očekáváme vliv analogů na podobu anglických slov v realizaci českých mluvčích. Pozornost je kladena především na temporální struktury slov a segmentů.

Druhým z výše zmíněných vybraných jevů je trvání v širším slova smyslu. Kromě jeho významu jako jednoho z projevů slovního přízvuku je trvání také inherentní součástí některých hlásek. Český fonologický inventář například obsahuje krátké a dlouhé samohlásky, které se od sebe odlišují pouze svou kvantitou. V angličtině se naopak všechny samohlásky odlišují i jinými parametry, a tak jejich trvání nemá tak výraznou rozlišovací funkci. Angličtina se od češtiny odlišuje také ve vlivu znělosti a neznělosti souhlásky na trvání předcházející samohlásky. Tyto rozdíly mezi oběma jazyky poskytují prostor pro zkoumání jejich interference.

V teoretické části se věnujeme rovněž poznatkům o angličtině českých mluvčích. V tomto ohledu práce navazuje na diplomovou práci Jana Růžka (2013), která se zabývá trváním hlásek ve víceslabičných slovech s českými analogy v angličtině českých mluvčích. Zatímco Růžek pracoval pouze se slovy, která mají analogy, tato práce zkoumá jak slova s analogy, tak slova bez nich, a porovnává jejich temporální struktury. Z dalších výzkumů české angličtiny jsou pro nás nejdůležitější práce zabývající se slovním přízvukem a trváním jako jedním z jeho projevů. Volín a Weingartová (2014) zjistili, že na rozdíl od základní frekvence nebo intenzity čeští mluvčí realizují temporální strukturu slov podobně jako rodilí mluvčí. Naopak jiné studie (Skarnitzl, 2005; Volín a Skarnitzl, 2010) dokazují, že i odlišnosti v trvání mají vliv na vznik a detekci cizineckého přízvuku v české produkci angličtiny.

Praktická část práce obsahuje samotný výzkum vlivu analogů na temporální strukturu anglických slov v produkci českých mluvčích. Naší hypotézou je, že slova s analogy budou temporálně odlišná od realizací rodilé mluvčí, zatímco slova, která analogy nemají, budou z hlediska trvání slov a segmentů imitována přesně. Dále také předpokládáme, že pevný slovní přízvuk v češtině bude mít za následek přesun temporální prominence na první slabiku u slov, která mají v angličtině přízvuk na druhé slabice.

V kapitole *Material and method* je popsán průběh výzkumu. Nejprve jsme pořídili nahrávky 36 anglických dvouslabičných slov v realizaci britské rodilé mluvčí. Tyto nahrávky pak byly přehrány 16 rodilým mluvčím češtiny, kteří měli za úkol co nejpřesněji imitovat podobu anglických slov. U všech českých mluvčích jsme provedli 4 opakování každého slova. Provedli jsme segmentaci pořízených nahrávek a získané hodnoty pro české mluvčí jsme následně porovnávali s příslušnými hodnotami v realizaci britské mluvčí.

Kapitola *Research part* se pak podrobněji věnuje analýze výsledků výzkumu a diskuzi. Podkapitola *Results* uvádí rozdíly v trvání mezi českou a nativní realizací anglických slov společně s jejich statistickou významností. Tyto rozdíly jsou zkoumány jak z hlediska přítomnosti analogů v českém jazyce, tak i z hlediska typu jednotky, kterou zkoumáme (slova, slabiky, segmenty a vybrané sekvence). Výsledky ukazují, že hlavním znakem imitovaných anglických slov v porovnání s nativní realizací je jejich obecně kratší trvání. Kromě toho se v této podkapitole zabýváme také rozdíly mezi jednotlivými mluvčími a mezi jednotlivými opakováními.

V podkapitole *Discussion* se k výsledkům vracíme a analyzujeme je s ohledem na cíl práce. Výsledky ukazují částečné potvrzení naší hypotézy. Porovnaná trvání celých slov i jednotlivých slabik ukazují, že u slov s analogy se rozdíly mezi realizacemi českých mluvčích a britské rodilé mluvčí prohlubují. Naopak ve slovech, která analogy nemají, se českým mluvčím daří temporální struktury slov a slabik realizovat přesněji. Zároveň se potvrdil předpoklad vlivu pevného přízvuku v češtině na trvání jednotlivých slabik. Z výsledků je patrné prodlužování první slabiky a zkracování druhé slabiky u anglických slov, a to ve slovech s přízvukem jak na první, tak na druhé slabice. V diskuzi zároveň předkládáme návrhy pro další studie, které by mohly ještě více prohloubit poznatky o zkoumané tématice.